IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of: PITMAN et al.

Serial No.: 09/275,568 Group Art Unit: 1631
Filed: 03/24/99 Examiner: Cheyne D. Ly

Title: Similarity Searching of Molecules Based Upon Descriptor Vectors Characterizing Molecular Regions

Commissioner for Patents
Mail Box Appeal Briefs
P.O. Box 1450
Alexandria, VA 22313-1450

CORRECTED APPEAL BRIEF

Sir:

This Appeal Brief is submitted in support of the Notice of Appeal filed September 29, 2005 and in response to the Notification of Non-Compliant Brief mailed on March 23, 2007.

REAL PARTY IN INTEREST

International Business Machines Corporation is the real party in interest as assignee of the subject application.

RELATED APPEALS AND INTERFERENCES

The Appellant, the Appellant's legal representative, and the Assignee are not aware of any other appeals or interferences which will directly affect, be directly affected by, or have a bearing on the Board's decision in this Appeal.

STATUS OF CLAIMS

Claims 1, 4-15 and 31-33 (the claims at issue) are pending in the above-identified patent application. Claims 2-3, 16-30 and 34-35 have been canceled. The claims at issue were finally rejected in an Office Action dated June 29, 2005. The final rejection of the claims at issue is hereby appealed.

The claims at issue all stand rejected under 35 U.S.C. § 101 as being directed to non-statutory subject matter. Claim 1 has been rejected under 35 U.S.C. § 102(e)(2) as being anticipated by Platt et al (U.S. Patent 5,784,294, hereafter "Platt"). The final Office Action rejected claims 1 and 4, under 35 U.S.C. § 112, first paragraph, as containing subject matter that was not described in the original specification so as to convey to one skilled in the art that the inventor possessed the claimed invention.

STATUS OF AMENDMENTS

A proposed amendment canceling claims 34-35 was filed with the Appeal Brief on July 21, 2006 and has been entered.

SUMMARY OF CLAIMED SUBJECT MATTER

The present invention, as set forth in claim 1, and as described and shown in the specification and the Figures of the above-identified patent application, is directed to a method for generating and storing data characterizing at least one region of said plurality of regions, the method comprises the steps of:

generating an entry [page 3, line 18] comprising i) an identifier that identifies said at least one region, and ii) data characterizing a set of axes derived from a property distribution of said at least one region [page 3, lines 19-20];

applying a mapping to the descriptor vector associated with said at least one region [page 3, lines 20-21] based on preselected criteria [page 3, lines 10-13];

generating a key that corresponds to said mapping of the descriptor vector associated with said at least one region [page 3, lines 20-21]; and

storing said entry in a memory [page 3, line 21], wherein said key is associated with said entry such that the key indexes the entry for retrieval thereof [page 4, lines 2-3].

A concept underlying the claimed invention is the storage of data in groupings that are sensitive to the way a human would search for stored information, thus facilitating retrieval of the stored data in a way that is useful for using the molecules.

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

I. Whether the Examiner erred in rejecting claims 1 and 4-15 and 31-33 as being directed to non-statutory subject matter.

II. Whether the Examiner erred in rejecting claim 1 under 35 U.S.C.§102(e)(2) as being anticipated by U.S. Patent Number 5,784,294 (Platt).

III. Whether the Examiner erred in rejecting claims 1 and 4 under 35 U.S.C. §112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the art that the inventor possessed the invention.

IV. Whether the Examiner erred in rejecting Claims 1 and 4-15 under 35 U.S.C. §112, second paragraph, for failing to point out and distinctly claim the subject matter which applicant regards as the invention.

GROUPING OF CLAIMS

For purposes of this Appeal, the claims at issue stand or fall together.

ARGUMENT

I. CLAIM REJECTIONS UNDER 35 U.S.C. § 101

The Examiner erred in rejecting the claims at issue under 35 U.S.C. §101 on grounds that the claimed invention is allegedly directed to non-statutory subject matter.

The analysis of whether a claim is directed to statutory subject matter begins with the language of 35 U.S.C. 101, which reads:

"Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title."

In AT&T Corp. v. Excel Communications, Inc., 172 F.3d 1352, 50 USPQ2d 1447 (Fed. Cir. 1999), the United States Court of Appeals for the Federal Circuit said that the Supreme Court has construed 101 broadly, noting that Congress intended statutory subject matter to "include anything under the sun that is made by man." See Diamond v. Chakrabarty, 447 U.S. 303, 309 (1980) (quoting S. Rep. No. 82-1979, at 5 (1952); H.R. Rep. No. 82-1923, at 6 (1952)); see also Diamond v. Diehr, 450 U.S. 175, 182 (1981). Notwithstanding the broad scope statutory subject matter, the Court has specifically identified three categories of unpatentable subject matter: "laws of nature, natural phenomena, and abstract ideas." See Diehr, 450 U.S. at 185.

In this appeal, all of the claims at issue are method claims which fall within the "process" category of the four enumerated categories of patentable subject matter

in 101. The examiner determined that the claims at issue recite a method that "is merely arranging the data based on an algorithm without any practical application."

The subject patent application claims methods or processes for generating and storing data. Moreover, these processes are performed by a machine (a data processing system). The data are expressed and processed as electrical signals operated upon by a processing apparatus. That is a practical application of the invention. The claim also states in the storing step that the key is associated with the subject data entry for retrieval thereof.

The Examiner's determination constitutes an error of law. The issue of failure to claim statutory subject matter is one of law that is reviewed *de novo*. See AT&T Corp. v. Excel Communications, Inc., supra.

The storage of data in a computer memory is by itself a concrete and useful result. Claim 1 which is representative includes four steps that precisely set forth how the subject information is stored in memory. In In re Lowry, 32 F. 3d 1579, 32 USPQ2d 1031 (Fed. Cir. 1994), the Federal Circuit held that claims directed to an invention related to "storage, use, and management of information residing in a memory were entitled to patentable weight." In Lowry, the Board of Patent

Appeals and Interferences reversed the statutory subject matter rejections made by the Examiner under 35 U.S.C. §101.

The Federal Circuit then went further by holding that claim limitations relating to the storage of information in a computer memory are entitled to patentable weight in distinguishing the prior art. The method claimed in the instant application is similar to the storage of data in Lowry. Storage of information in a computer memory is an important aspect of information technologies because information processing apparatus must read data from computer memory to execute operations on that information. Considering the claims at issue as a whole, as they must be, it becomes clear that the information stored in a memory has a practical purpose beyond the mere storage of the information – retrieval of the stored information based on the key mapped to a descriptor vector. The ability to retrieve information from memory based on various criteria is perhaps as important as storage of the information.

The technology of search engines which is the subject of numerous patents is concerned with this very concept. Failure to provide patent protection to inventions in the art of data retrieval would violate the constitutional mandate of promoting the progress of science and the useful arts. Therefore, the rejection of the claims at issue for failure to recite statutory subject matter must be reversed.

In the final Office Action dated January 27, 2003, the Examiner rejected Appellants' arguments and reasserted the position that the claims are directed to non-patentable subject matter. The Examiner thus contended that: "An invention where a system merely stores data such as descriptor vectors associated with a plurality of regions of molecules onto a media [sic, medium] is considered to be non-statutory subject matter because the said data is considered to be nonfunctional descriptive material." This rejection is nothing more than an application of the "printed matter" category of non-patentable subject matter that was so clearly discredited and reversed in the In re Lowry decision. As noted above, the general rule of patentable subject matter is expansive and any determination of failure to claim statutory subject matter must find its support in the case law.

In the final Office Action the Examiner concedes that the claimed invention does not lack utility under section 101 of the patent statute. Therefore, the rejection is either based on the printed matter exception or on the algorithm exception to the rule of patentability. As noted above, storage of data (which is by its nature descriptive) is a very important aspect of the information technology arts. That fact was recognized in In re Lowry when the Federal Circuit laid to rest the doctrine of printed matter as applied to data stored in computer readable media. The Federal Circuit's decision in In re Lowry requires reversal of the Examiner's determination of failure to claim statutory subject matter.

To the extent that the Examiner's section 101 rejection relied on the mathematical algorithm doctrine it must also be reversed. In AT&T, supra, the Federal Circuit said that any step-by-step process involves an algorithm in the broad sense of the word. The AT&T court thus said: "Since the process of manipulation of numbers is a fundamental part of computer technology, we have had to reexamine the rules that govern the patentability of such technology. The sea-changes in both law and technology stand as a testament to the ability of law to adapt to new and innovative concepts, while remaining true to basic principles." AT&T, 172 F.3d at 1356. Thus the AT&T court limited the "Algorithm" doctrine to apply only in cases of purely "abstract" algorithms. See AT&T at 1357. In AT&T, the Federal Circuit also said that the algorithm must be applied in a useful way and found a practical result in the claimed methods in the addition of certain descriptive information called a PIC (or primary interchange carrier) to certain other information used in switching telephone calls. The information in the claims at issue in the instant case also has a useful result – the storage for retrieval of information from a computer memory responsive to a search for certain criteria. The retrieved information is useful for, among other purposes, determining properties of molecules.

The Examiner's argument that the information stored according to the claims at issue is merely descriptive if applied to the field of photography would preclude

the patentability of cameras because cameras take light, one form of information that represents an object, and record the information in film. The information is merely descriptive of the subject of the photograph. The application of the Examiner's reasoning to the clearly patentable area of photography illustrates the point that the claimed invention, which is analogous to other forms of data storage, should not be precluded from patentability.

The Federal Circuit rejected an argument similar to the Examiner's Arrhythmia Research Technology, Inc. v. Corazonix Corp., 22 USPQ2d 1033 (Fed. Cir. 1992), where processing information describing a patient's heartbeat was held to be statutory subject matter. The court there said that the claims at issue did not preempt all uses of the algorithm, Arrhythmia at 1060. Similarly, in the instant case the claims do not preempt all uses of any algorithms; rather they are limited to storage and retrieval in a computer memory. Therefore, Appellants request reversal of the rejection of the claims at issue under 35 U.S.C. §101.

II. CLAIM REJECTIONS UNDER 35 U.S.C. §102(e)(2)

Claim 1 was rejected under 35 U.S.C. § 102(e)(2) as being anticipated by Platt (U.S. Pat. No. 5,784,294). This rejection should be reversed because the Examiner has not shown that claim 1 is anticipated by Platt. Nowhere does Platt teach or disclose *any* of the elements of claim 1. Platt relates to a storage device that performs a plurality of functions that produce a result that can be an input to

the method of claim 1 of the instant application but which does not anticipate the claims at issue. Platt does not disclose the required mapping, generation of a key, or storing the entry as required by claim 1.

In the final Office Action, the Examiner contends that Platt discloses at Fig. 9 "storing an entry comprising a molecular descriptor with a key to access it." Fig. 9 of Platt is a flow chart illustrating use of descriptors. It does not relate to a mapping of descriptor vectors (as claimed) at all. The key generated according to claim 1 corresponds to the mapping. The Examiner has not shown how Fig. 9 of Platt performs a mapping or any of the claimed steps.

The Examiner further says "Platt et al. teaches storing said first and second descriptors of each molecule in said series of molecules in a database for subsequent processing to thereby identify correspondence between molecules in said series of molecules (Claim 34, Lines 39-42)." That statement does not describe <u>any</u> of the claimed steps. The claimed step of "storing" relates to the entry defined in the first step. The section of Pratt cited has nothing to do with such an entry and hence cannot correspond to the claimed storing step, or any other claimed step.

The Examiner also argues that the "key" is inherent. Again, the claimed "key" corresponds to the claimed mapping and the Examiner has not shown anything in

Platt corresponding to such a mapping. Instead, the Examiner argues that Platt has some criteria for selecting molecules of the training set to be placed in a table and that this corresponds to "applying the mapping." The Examiner does not show how the placement of molecules in a table relates even remotely to applying a mapping to a descriptor vector as claimed and has fallen far short of the exact relationship that anticipation requires.

Finally, the Examiner has erred as a matter of law by arguing that a type of data structure allegedly disclosed in Platt "is consistent with" the limitation of "key indexes to entry for retrieval thereof." The legal test for anticipation is whether every element of a claim is found in an item of prior art, and not whether a structure is consistent with a claimed method. Therefore, Appellants request reversal of the rejections under Section 102(e).

III. CLAIM REJECTIONS UNDER 35 U.S.C. § 112, FIRST PARAGRAPH

The final Office Action rejected claims 1 and 4 under 35 U.S.C. § 112, first paragraph, as containing subject matter that was not described in the original specification so as to convey to one skilled in the art that the inventor possessed the claimed invention. The Examiner (or the Board, if the Board is the first body to raise a particular ground for rejection) "bears the initial burden . . . of presenting a *prima facie* case of unpatentability. "In re Oetiker, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). Insofar as the written description

requirement is concerned, that burden is discharged by "presenting evidence or reasons why persons skilled in the art would not recognize in the disclosure a description of the invention defined by the claims." In re Wertheim, 541 F.2d 257, 263, 191 USPQ 90, 97 (CCPA 1976). Thus, the burden placed on the Examiner varies, depending upon what the applicant claims. The specification contains a description of the claimed invention, albeit not in *ipsis verbis* (in the identical words). Then the Examiner or Board, in order to meet the burden of proof, must provide reasons why one of ordinary skill in the art would not consider the description sufficient. Id. at 264, 191 USPQ at 98. In the present case, the amendment of November 13, 2002 amended claim 1 so that the step of applying a mapping to the descriptor vector is based on pre-selected criteria. Support for the amendment does not have to be *ipsis verbis*. It is inherent from the discussion in page 40 of the specification that the application of the mapping is based on pre-determined criteria. Note that the discussion (page 40) of the "association criteria" is defined in the prior training phase and thus clearly the association criteria were "pre-determined."

Claim 1 was further amended to state that "the key indexes the entry for retrieval thereof." It is inherent in the claimed invention that the key indexes the entry for retrieval of the entry. Why else would a key corresponding to a mapping be used? In any case, the gist of the written description requirement is to prevent an applicant from adding claims to subject matter that the inventor did not possess at

the time of filing. Vas-Cath Inc., v. Mahurkar, 935 F.2d 1555, 19 USPQ2d 111 (Fed. Cir. 1991). As Appellants noted in the amendment of November 13, 2002, the amendment was not made to define additional subject matter but to make clear what was already implicit. Appellant again poses the question – why would information be stored if not for retrieval thereof?

The Examiner has not shown any reason why the language added in the amendment would not be supported by the specification and in fact Appellants contend that the amendment was not made for purposes of patentability, so the invention defined by the claims both before and after the amendment are the same and hence was clearly in the possession of the inventor at the time of the filing of the application. Therefore, Appellants request reversal of this rejection.

IV. CLAIM REJECTIONS UNDER 35 U.S.C. §112, SECOND PARAGRAPH

The final Office Action also rejected claims 1 and 4-15 under 35 U.S.C. §112, second paragraph, as being indefinite. Specifically, the final Office Action contends that claim 1 is vague and indefinite due to lack of clarity in the preamble and that it is not clear to the Examiner whether the Appellant intended to claim a data processing system or a method. Appellant contends that this ground for rejection was in error and requests reversal thereof. Claim 1 clearly states that it is directed to a method for generating and storing data and recites a series of steps. The reference to a data processing system is preceded by the word "in"

and is thus an introductory phrase that merely indicates a field of use for the claimed method. See Kropa v. Robie, 187 F.2d 150, 158 (CCPA 1951).

The final Office Action has not shown that claim 1 is indefinite because it would be an error of law to construe claim 1 as being directed to a system. Claims 4-15 were rejected on the basis that they are dependent on claim 1 and their rejection should also be reversed on the foregoing grounds.

CONCLUSION

In view of the foregoing, it is respectfully submitted that the application and the claims are in condition for allowance. Reversal of the final rejection, and allowance of the claims as amended, are requested.

Respectfully submitted,

By: Michael J. Buchenhorner   Registration Number 33,162

Michael J. Buchenhorner
Attorney for Appellants

E-Filed on Date: **April 23, 2007**

CLAIMS APPENDIX

1. In a data processing system wherein descriptor vectors associated with a plurality of regions of molecules are stored in a database, a method for generating and storing data characterizing at least one region of said plurality of regions, the method comprising the steps of:

generating an entry comprising i) an identifier that identifies said at least one region, and ii) data characterizing a set of axes derived from a property distribution of said at least one region;

applying a mapping to the descriptor vector associated with said at least one region based on preselected criteria;

generating a key that corresponds to said mapping of the descriptor vector associated with said at least one region; and

storing said entry in a memory, wherein said key is associated with said entry such that the key indexes the entry for retrieval thereof.

2. The method of claim 1, wherein said set of axes are invariant to rotation and translation of said at least one region.

3. The method of claim 2, wherein said set of axes are derived from principal axes of said property distribution.

4. The method of claim 1, wherein said property distribution of said at least one region is computed from a convolution with a probe function to a property field.

5. The method of claim 1, wherein said plurality of descriptor vectors are classified into groups, and wherein said mapping step maps said descriptor vectors to a space discriminating between said groups of descriptor vectors.

6. The method of claim 5, wherein said mapping is derived from the steps of:

generating first data representing differences between said groups of descriptor vectors;

generating second data representing variations within said groups of descriptor vectors;

identifying a set of component vectors that maximizes an F distributed criterion function, said criterion function having a numerator based upon said first data and a denominator based upon said second data;

generating an F distributed statistic for subsets of said component vectors, said statistic having a numerator based upon said first data and a denominator based upon said second data;

for each particular subset of component vectors, calculating a probability value for the F-distributed statistic associated with the particular subset;

selecting a probability value from probability values for said subsets of component vectors based upon a predetermined criterion;

identifying the subset of said component vectors associated with the selected probability value; and

generating a mapping to a space corresponding to the subset of component vectors associated with the selected probability value, and storing the mapping for subsequent processing.

7. The method of claim 6, wherein said first data comprises a matrix $\ddot{y}_b$ representing covariance between said groups of descriptor vectors, and said second data comprises a matrix $\ddot{y}_W$ representing covariance within said groups of descriptor vectors.

8. The method of claim 7, wherein said criterion function has the general form:

$$f(\hat{w}) = C\left(\frac{\hat{w}^T \varepsilon_b \hat{w}}{\hat{w}^T \varepsilon_w \hat{w}}\right)$$

where $\hat{w}$ is some vector, T indicates a transpose, ε_b is a first data representing covariance, ε_w is a second data representing covariance and C is a constant based upon degrees of freedom in ε_b and ε_w.

9. The method of claim 8, wherein C is determined as follows:

$$C = \frac{1/\text{degrees of freedom in } \varepsilon_b}{1/\text{degrees of freedom in } \varepsilon_w} = \frac{1/(N-1)}{1/(\Sigma\ n_i - N)}$$

where N represents the number of groups of descriptor vectors, n_i represents the number of regions, and $\Sigma\ n_i$ represents the sum of n_i for the N groups.

10. The method of claim 7, wherein the step of identifying a set of component vectors that maximizes an F distributed criterion function comprises the substeps of:

determining a set of (eigenvalue, eigenvector) pairs for the matrix $ÿ_W$

determining said set of component vectors based upon said set of (eigenvalue, eigenvector) pairs for the matrix $ÿ_W$.

11. The method of claim 10, wherein said statistic for a given subset of component vectors is based upon value of said criterion function for said subset of component vectors.

12. The method of claim 11, wherein said statistic for a given subset of component vectors has the following form:

$$\Psi_s = C\ (1/L_S)\ \Sigma f_k$$

where f_k represents the value of the criterion function at a component vector in the given subset,

C is a constant,

represents the number of values in the given subset of component vectors, and the Σ operation sums over the L_S f_k values in the given subset of component vectors.

13. The method of claim 12, wherein said a probability value for a particular F-distributed statistic represents a probability value that the particular F-distributed statistic could have been larger by chance.

14. The method of claim 13, wherein said probability value selected from probability values for said subsets of component vectors is a minimum probability value of said probability values for said subsets of component vectors.

15. The method of claim 6, wherein said mapping for said at least one descriptor vector performs a loop over each component vector belonging to the subset of component vectors associated with the selected probability;

wherein, in each iteration of said loop, dot product of said descriptor vector with a transpose of a unit vector for the given component vector is added to a running sum.

31. The method of claim 1, wherein the at least one descriptor vector is invariant to rotation and translation of the at least one region.

32. The method of claim 31, wherein the set of axes is derived from principal axes of second moments of a region of the property distribution information.

33. The method of claim 6, wherein the probability value is obtained by treating the ratio as an F-distributed statistic.

EVIDENCE APPENDIX

STRUCTURAL MEDICINE **Protein Crystallography Course**

The convolution theorem and its applications

Outline

What is a convolution?

One of the most important concepts in Fourier theory, and in crystallography, is that of a convolution. Convolutions arise in many guises, as will be shown below. Because of a mathematical property of the Fourier transform, referred to as the convolution theorem, it is convenient to carry out calculations involving convolutions.

But first we should define what a convolution is. Understanding the concept of a convolution operation is more important than understanding a proof of the convolution theorem, but it may be more difficult!

Mathematically, a convolution is defined as the integral over all space of one function at x times another function at u-x. The integration is taken over the variable x (which may be a 1D or 3D variable),

typically from minus infinity to infinity over all the dimensions. So the convolution is a function of a new variable u, as shown in the following equations. The cross in a circle is used to indicate the convolution operation.

$$C(\mathbf{u}) = f(\mathbf{x}) \otimes g(\mathbf{x}) = \int_{space} f(\mathbf{x})\, g(\mathbf{u}-\mathbf{x})\, d\mathbf{x}$$

$$= g(\mathbf{x}) \otimes f(\mathbf{x}) = \int_{space} g(\mathbf{x})\, f(\mathbf{u}-\mathbf{x})\, d\mathbf{x}$$

Note that it doesn't matter which function you take first, *i.e.* the convolution operation is commutative. We'll prove that below, but you should think about this in terms of the illustration below. This illustration shows how you can think about the convolution, as giving a weighted sum of shifted copies of one function: the weights are given by the function value of the second function at the shift vector. The top pair of graphs shows the original functions. The next three pairs of graphs show (on the left) the function g shifted by various values of x and, on the right, that shifted function g multiplied by f at the value of x.

$$f(x) \otimes g(x) = g(x) \otimes f(x)$$

We stated that the convolution integral is commutative. Here, in case you are interested, is a quick proof of that. First, we start with the convolution integral written one way. For convenience we will deal with the 1D case, but the 3D case is exactly analogous.

$$C(u) = \int_{-\infty}^{\infty} f(x) g(u-x) dx$$

Now we substitute variables, replacing u-x with a new x'.

$$x' = u - x, \; dx' = -dx$$

$$C(u) = -\int_{\infty}^{-\infty} f(u-x') g(x') dx'$$

Note that, because the sign of the variable of integration changed, we have to change the signs of the limits of integration. Because these limits are infinite, the shift of the origin (by the vector u) doesn't change the magnitude of the limits.

Now we reverse the order of the limits, which changes the sign of the equation, and swap the order of the functions g and f.

$$C(u) = \int_{-\infty}^{\infty} g(x') f(u-x') dx'$$

It doesn't matter whether we call the variable of integration x' or x, so we put back x, to get the result we wanted to prove.

$$C(u) = \int_{-\infty}^{\infty} g(x) f(u-x) dx$$

The convolution theorem

Because there will be so many Fourier transforms in the rest of this presentation, it is useful to introduce a shorthand notation. T will be used to indicate a forward Fourier transform, and its inverse to indicate the inverse Fourier transform.

$$T(f(r)) = \int_{space} f(r)\exp(2\pi i s \cdot r)dr$$

There are two ways of expressing the convolution theorem:

- The Fourier transform of a convolution is the product of the Fourier transforms.
- The Fourier tranform of a product is the convolution of the Fourier transforms.

$$T(f \otimes g) = T(f)T(g)$$
$$T(f\,g) = T(f) \otimes T(g)$$

The convolution theorem is useful, in part, because it gives us a way to simplify many calculations. Convolutions can be very difficult to calculate directly, but are often much easier to calculate using Fourier transforms and multiplication.

Proof of the convolution theorem

To prove the convolution theorem, in one of its statements, we start by taking the Fourier transform of a convolution. What we want to show is that this is equivalent to the product of the two individual Fourier transforms. Note, in the equation below, that the convolution integral is taken over the variable x to give a function of u. The Fourier transform then involves an integral over the variable u.

$$\begin{aligned} T(f(x) \otimes g(x)) &= T\left(\int_{-\infty}^{\infty} f(x)g(u-x)dx\right) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)g(u-x)dx \exp(2\pi i s u)du \end{aligned}$$

Now we substitute a new variable w for u-x. As above, the infinite integration limits don't change. Then we expand the exponential of a sum into the product of exponentials and rearrange to bring together expressions in x and expressions in w.

$$\begin{aligned} T(f(x) \otimes g(x)) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)g(w)\exp[2\pi i s(x+w)]dx\,dw \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)\exp(2\pi i s x)g(w)\exp(2\pi i s w)dx\,dw \end{aligned}$$

Expressions in x can be taken out of the integral over w so that we have two separate integrals, one over x with no terms containing w and one over w with no terms containing x.

$$T(f(x)\otimes g(x)) = \int_{-\infty}^{\infty} f(x)\exp(2\pi isx)dx \int_{-\infty}^{\infty} g(w)\exp(2\pi isw)\,dw$$

The variables of integration can have any names we please, so we can now replace w with x, and we have the result we wanted to prove.

$$\begin{aligned} T(f(x)\otimes g(x)) &= \int_{-\infty}^{\infty} f(x)\exp(2\pi isx)dx \int_{-\infty}^{\infty} g(x)\exp(2\pi isx)\,dx \\ &= T(f(x))\,T(g(x)) \end{aligned}$$

If you look through the derivation above, you will see that we could have used a minus sign in the exponential when taking the original Fourier transform, and then the two Fourier transforms at the end would also contain minus signs in the exponentials. In other words, the convolution theorem applies to both the forward and reverse Fourier transforms. This is not surprising, since the two directions of Fourier transform are essentially identical.

Proof of second statement of convolution theorem

To prove the second statement of the convolution theorem, we start with the version we have already proved, *i.e.* that the Fourier transform of a convolution is the product of the individual Fourier transforms.

$$T(f\otimes g) = T(f)\,T(g)$$

First we'll define some shorthand, where capital letters indicate the Fourier transform mates of lower case letters.

$$\begin{aligned} F &= T(f) \quad & f &= T^{-1}(F) \\ G &= T(g) \quad & g &= T^{-1}(G) \end{aligned}$$

We use these relationships to recast the statement above in terms of the Fourier tranform mates of the original functions. Then we take an inverse Fourier transform on each side of the equation, to get (essentially) the second statement of the convolution theorem. The only difference is that it is expressed in terms of the inverse Fourier transform.

$$T\left[T^{-1}(F)\otimes T^{-1}(G)\right]=F\,G$$
$$T^{-1}(F)\otimes T^{-1}(G)=T^{-1}(F\,G)$$

But, as we noted above, we could have proved the convolution theorem for the inverse transform in the same way, so we can reexpress this result in terms of the forward transform.

$$T(f\,g)=T(f)\otimes T(g)$$

The correlation theorem

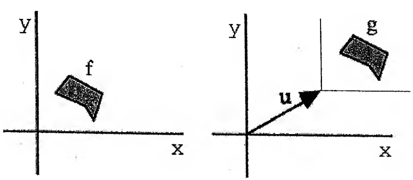
The correlation theorem is closely related to the convolution theorem, and it also turns out to be useful in many computations. The correlation theorem is a result that applies to the correlation function, which is an integral that has a definition reminiscent of the convolution integral.

What is a correlation function?

In a correlation integral, instead of taking the value of one function at u-x, you take the value of that function at x-u. Equivalently, you take the value of the other function at x+u. This is shown in the following equation, along with the variable substitution that allows the two expressions to be interconverted.

$$f\circ g=\int_{-\infty}^{\infty} f(x)\,g(x+u)\,dx;\ \text{substitute}\ x'=x+u$$
$$=\int_{-\infty}^{\infty} f(x'-u)\,g(x')\,dx'$$
$$=\int_{-\infty}^{\infty} f(x-u)\,g(x)\,dx$$

The figure below illustrates why the correlation function has the name that it does. If the two functions f and g contain similar features, but at a different position, the correlation function will have a large value at a vector corresponding to the shift in the position of the feature.



The correlation theorem can be stated in words as follows: the Fourier tranform of a correlation integral is equal to the product of the complex conjugate of the Fourier transform of the first function and the Fourier transform of the second function. The only difference with the convolution theorem is in the presence of a complex conjugate, which reverses the phase and corresponds to the inversion of the argument u-x.

$$T(f \circ g) = \mathbf{F^* \, G}$$

Parseval's theorem

Important convolutions

Convolution with a Gaussian

First we need to define a Gaussian function. We will stick, for the moment, to 1D Gaussians. The Gaussian function is the familiar bell-shaped curve, with a peak position (r_0) and standard deviation.

$$p(r) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r-r_0)^2}{2\sigma^2} \right)$$

We won't derive the Fourier transform of a Gaussian, but it is given by the following equation.

$$T(p(r)) = \exp(2\pi i \, r_0 s) \exp(-2\pi^2 \sigma^2 s^2)$$

Note that the Fourier transform of a Gaussian is another Gaussian (although lacking the normalisation constant). There is a phase term, corresponding to the position of the center of the Gaussian, and then the negative squared term in an exponential. Also notice that the standard deviation has moved from the denominator to the numerator. This means that, as a Gaussian in real space gets broader, the corresponding Gaussian in reciprocal space gets narrower, and vice versa. This makes sense, if you think about it: as the Gaussian in real space gets broader, contributions from points within that Gaussian start

to interfere with each other at lower and lower resolutions.

Convolution with a Gaussian will shift the origin of the function to the position of the peak of the Gaussian, and the function will be smeared out, as illustrated above.

Convolution with a delta function

Delta functions have a special role in Fourier theory, so it's worth spending some time getting acquainted with them. A delta function is defined as being zero everywhere but for a single point, where it has a weight of unity.

$$\delta(\mathbf{r}-\mathbf{r}_0) = \begin{matrix} \text{weight of 1 at } \mathbf{r}=\mathbf{r}_0 \\ 0 \text{ elsewhere} \end{matrix}$$

What it means to say that it has a weight of unity is that the integral of the delta function over all space is 1. The delta function is given an argument of $\mathbf{r}$-$\mathbf{r}_0$ so that it can be defined as having its non-zero point at the origin. When $\mathbf{r}$ is equal to $\mathbf{r}_0$, the argument of the delta function is zero.

$$\int_{space} \delta(\mathbf{r}-\mathbf{r}_0)d\mathbf{r} = 1$$

A more general property of the delta function is that the integral of a delta function times some other function is equal to the value of that other function at the position of the delta function.

$$\int_{space} \mathrm{f}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}_0)d\mathbf{r} = \mathrm{f}(\mathbf{r}_0)$$

How can a single point, with no width, breadth or depth, have a weight of one? The value of the delta function at that point must be a special kind of infinity, and this means that it has to be defined as a limit. There are a number of ways to define a delta function. One of them is to define it as an infinitely sharp Gaussian. The integral over all space of a Gaussian is 1, which satisfies one of the properties required for the delta function, and if we take the limit of a Gaussian as the standard deviation tends to zero, it satisfies the other properties. The following equation defines a 3D delta function as the limit of an isotropic 3D Gaussian.

$$\delta(\mathbf{r}-\mathbf{r}_0) = \lim_{\sigma^2 \to 0} \frac{1}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{|\mathbf{r}-\mathbf{r}_0|^2}{2\sigma^2}\right)$$

With this definition of the delta function, we can use the Fourier transform of a Gaussian to determine the Fourier transform of a delta function. As the standard deviation of a Gaussian tends to zero, its Fourier transform tends to have a constant magnitude of 1. All that is left is the phase shift term.

$$T\left[\delta(\mathbf{r}-\mathbf{r}_0)\right]=\exp(2\pi i\mathbf{s}\cdot\mathbf{r}_0)$$

So we see that the Fourier transform of a delta function is just a phase term. Think about the picture we had of an electron at a point; it contributed just a phase term, with unit weight, to the diffraction pattern. So now we see that we can consider an electron at a point to be a delta function of electron density.

Finally we can consider the meaning of the convolution of a function with a delta function. If we write down the equation for this convolution, and bear in mind the property of integrals involving the delta function, we see that convolution with a delta function simply shifts the origin of a function.

$$\int_{space}\delta(\mathbf{r}-\mathbf{r}_0)\mathrm{f}(\mathbf{u}-\mathbf{r})d\mathbf{r}=\mathrm{f}(\mathbf{u}-\mathbf{r}_0)$$

$$\delta(\mathbf{r}-\mathbf{r}_0)\otimes\mathrm{f}(\mathbf{r})=\mathrm{f}(\mathbf{r}-\mathbf{r}_0)$$

Applications of the convolution theorem

Atomic scattering factors

We have essentially seen this before. We can tabulate atomic scattering factors by working out the diffraction pattern of different atoms placed at the origin. Then we can apply a phase shift to place the density at the position of the atom. Our new interpretation of this is that we are convoluting the atomic density distribution with a delta function at the position of the atom.

B-factors

We can think of thermal motion as smearing out the position of an atom, *i.e.* convoluting its density by some smearing function. The B-factors (or atomic displacement parameters, to be precise) correspond to a Gaussian smearing function. At resolutions typical of protein data, we are justified only in using a single parameter for thermal motion, which means that we assume the motion is isotropic, or equivalent in all directions. (In crystals that diffract to atomic resolution, more complicated models of thermal motion can be constructed, but we won't deal with them here.)

Above, we worked out the Fourier transform of a 1D Gaussian.

$$\mathrm{T}(\mathrm{p}(r))=\exp(2\pi i\, r_0 s)\exp\left(-2\pi^2\sigma^2 s^2\right)$$

In fact, all that matters is the displacement of the atom in the direction parallel to the diffraction vector, so this equation is suitable for a 3D Gaussian. All we have to remember is that the term corresponding to the standard deviation refers only to the direction parallel to the diffraction vector. Since we are dealing with the isotropic case, the standard deviation (or atomic displacement) is equal in all directions.

The B-factor is used in an equation in terms of sinθ/λ instead of the diffraction vector, because all that matters is the magnitude of the diffraction vector. We replace the variance (standard deviation squared) by the mean-square displacement of the atom in any particular direction. The B-factor can be defined in terms of the resulting equation.

$$|\mathrm{s}| = \frac{1}{d} = 2\sin\theta / \lambda$$

$$\mathrm{T}(\mathrm{p}(\mathrm{r})) = \exp\left(-8\pi^2 \left\langle u_x^2 \right\rangle \sin^2\theta / \lambda^2\right)$$

$$B = 8\pi^2 \left\langle u_x^2 \right\rangle$$

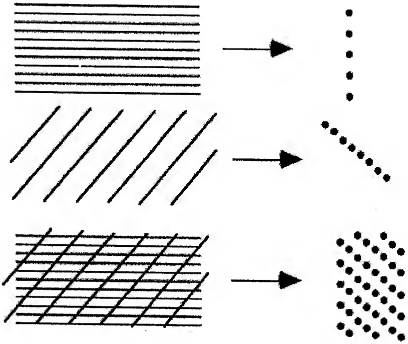
Note that there is a common source of misunderstanding here. The mean-square atomic displacement refers to displacement in any particular direction. This will be equal along orthogonal x, y and z axes. But often we think of the mean-square displacement as a radial measure, *i.e.* total distance from the mean position. The mean-square radial displacement will be the sum of the mean-square displacements along x, y and z; if these are equal it will be three times the mean-square displacement in any single direction. So the B-factor has a slightly different interpretation in terms of radial displacements.

$$\left\langle u_r^2 \right\rangle = \left\langle u_x^2 \right\rangle + \left\langle u_y^2 \right\rangle + \left\langle u_z^2 \right\rangle = 3\left\langle u_x^2 \right\rangle$$

$$B = \frac{8\pi^2}{3} \left\langle u_r^2 \right\rangle$$

Diffraction from a lattice

The convolution theorem can be used to explain why diffraction from a lattice gives



If you've gotten this far, I'm sorry that I ran out of time to complete this document before giving the lecture! The rest will be filled in at some point after all the lectures (and associated web pages) are finished.

Diffraction from a crystal

Resolution truncation

Density modification

Solvent flattening

Sayre's equation

Applications of the correlation theorem

The Patterson function

The phased translation function

Last updated: 7 June, 2005

The difference between this and Dijkstra's algorithm is that in choosing which vertex and edge to adjoin, we minimize a_{ij}, not d_j $(= d_i + a_{ij})$. We can implement Prim's algorithm to take about n^2 steps in a similar way, by using two arrays, one to tell us the current minimum weight on an edge from i to a point of S (only relevant for $i \notin S$), and another to tell us which previous vertex i is joined to; the value of the latter is fixed once $i \in S$.

Example 8.10

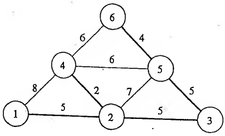


Figure 8.34

We apply Prim's algorithm to the edge-weighted graph shown in Figure 8.34. The edges given by the algorithm are shown in bold.

S	d_1	d_2	d_3	d_4	d_5	d_6	previous vertex to 1	2	3	4	5	6
1		5	∞	8	∞	∞		1		1		
1,2			5	2	7	∞		1	2	2	2	
1,2,4			5		6	6		1	2	2	4	4
1,2,3,4					5	6		1	2	2	3	4
1,2,3,4,5						4		1	2	2	3	5
1,2,3,4,5,6								1	2	2	3	5

Exercises 8.3

1. Show that a relation having adjacency matrix A is
   (a) reflexive if and only if $I + A = A$
   (b) symmetric if and only if $A^T = A$ (where A^T, the **transpose** of A, is obtained by interchanging its rows and columns)
   (c) transitive if and only if $A + A^2 = A$
   where 'addition' is taken as max on $\{0, 1\}$.

2. Find the transitive closures of the relations R_1 and R_2 on $\{0, 1, 2, \ldots, 20\}$ given by
   $(x, y) \in R_1$ if $y = x + 3$
   $(x, y) \in R_2$ if $y = x + 3 \bmod 21$
   Is either of R_1 and R_2 an equivalence relation? What about TC (R_1) or TC (R_2)?

RELATED PROCEEDINGS APPENDIX

None.